WHAT IS CLAIMED IS:

1. A method for ranking documents, comprising:

creating a ranking model that predicts a likelihood that a document will be selected;

training the ranking model using a data set that includes approximately tens of millions of instances;

identifying documents relating to a search query;

scoring the documents based, at least in part, on the ranking model; and

forming search results for the search query from the scored documents.

2. The method of claim 1, wherein the creating a ranking model includes:

storing information associated with a plurality of prior searches,

determining a prior probability of selection based, at least in part, on the information associated with the prior searches, and

generating the ranking model based, at least in part, on the prior probability of selection.

3. The method of claim 2, wherein the information associated with the prior searches includes, for each of a plurality of documents associated with the prior searches, a position occupied by the document within prior search results, a score assigned to the document, and a number of documents listed above the document in the prior search results that were selected.

4. The method of claim 1, wherein the creating a ranking model includes:

storing training data,

extracting features from the training data, and

generating conditions that include one or more of the extracted features.

5. The method of claim 4, wherein the creating a ranking model further includes:

selecting one of the conditions as a candidate condition,

estimating a weight for the candidate condition,

forming a new rule from the candidate condition and the estimated weight,

comparing a likelihood of the training data between a current model with the new rule and the current model without the new rule, and

selectively adding the new rule to the current model based, at least in part, on a result of the comparison.

6. The method of claim 5, wherein the selecting one of the conditions as a candidate condition includes at least one of:

creating the candidate condition from combinations of features or complements of features in the training data,

randomly selecting one of the conditions as the candidate condition,

selecting one of the conditions that includes a single one of the features as the candidate condition, and

augmenting one of the conditions by adding one or more features to the one condition to form the candidate condition.

7. The method of claim 5, wherein the estimating a weight includes determining a weight that maximizes a likelihood of the training data given the model.

8. The method of claim 5, wherein the selectively adding the new rule to the current model includes adding the new rule to the current model when the likelihood of the training data when the current model includes the new rule is sufficiently greater than when the current model does not include the new rule.

9. The method of claim 5, wherein the selectively adding the new rule to the current model further includes:

associating a cost with each of the conditions, and

determining whether to add the new rule to the current model based, at least in part, on the cost associated with the candidate condition.

10. The method of claim 5, further comprising:

performing a number of iterations including estimating the weight, forming the new rule, and comparing the likelihood of the training data.

11. The method of claim 1, wherein the data set also includes approximately millions of features.

12. The method of claim 1, wherein the scoring the documents includes:

forming an instance that corresponds to the search query and one of the documents, extracting features associated with the instance,

identifying rules in the ranking model that apply based, at least in part, on the extracted features, each of the identified rules including a weight, and

combining the weights of the identified rules with a prior probability of selection corresponding to the instance to generate a score for the one document.

13. The method of claim 12, wherein the instance includes user information corresponding to a user who provided the search query, query data corresponding to the search query, and document information corresponding to the one document.

14. The method of claim 1, wherein the scoring the documents includes:

determining a prior probability of selection corresponding to the search query and one of the documents, and

generating a score for the one document based, at least in part, on the determined prior probability of selection.

15. The method of claim 14, wherein the generating a score for the one document includes using the determined prior probability of selection as one of a plurality of factors in determining the score for the one document.

16. A system for ranking documents, comprising:

means for receiving a search query;

means for identifying documents relating to the search query;

means for ranking the documents based, at least in part, on a ranking model trained on a large data set that includes approximately millions of features; and

means for forming search results for the search query from the ranked documents.

17. A system for ranking documents, comprising:

a repository configured to store information corresponding to a plurality of prior searches; and

a server configured to:

receive a search query from a user,

identify documents corresponding to the search query, and

rank the identified documents based, at least in part, on a ranking model that includes rules that maximize a likelihood of the repository.

18. The system of claim 17, wherein the information in the repository includes a plurality of instances that include user information, query data, and document information corresponding to the plurality of prior searches.

19. The system of claim 18, wherein the repository is further configured to store a plurality of features associated with the instances.

20. The system of claim 19, wherein when ranking the documents, the server is configured to:

identify one of the instances that corresponds to the search query and one of the identified documents,

determine features associated with the identified instance,

identify rules in the ranking model that apply based, at least in part, on the determined features, each of the identified rules including a weight, and

combine the weights of the identified rules with a prior probability of selection corresponding to the identified instance to determine a rank for the one document.

21. The system of claim 20, wherein the user information of the identified instance includes information corresponding to the user who provided the search query, the query data of the identified instance includes information corresponding to the search query, and the document information of the identified instance includes information corresponding to the one document.

22. The system of claim 17, wherein when ranking the documents, the server is configured to:

determine a prior probability of selection corresponding to the search query and one of the identified documents, and

determine a rank for the one document based, at least in part, on the determined prior probability of selection.

23. The system of claim 22, wherein when determining a rank for the one document, the server is configured to use the determined prior probability of selection as one of a plurality of factors in determining the rank for the one document.

24. The system of claim 17, wherein the repository stores approximately tens of millions of instances and approximately millions of features associated with the plurality of prior searches.

25. A system for generating a model, comprising:

a repository configured to store training data that includes a plurality of features; and

one or more devices configured to:

select a candidate condition that includes one or more of the features,

estimate a weight for the candidate condition,

form a new rule from the candidate condition and the weight,

compare a likelihood of the training data between a model with the new rule and the model without the new rule, and

selectively add the new rule to the model based, at least in part, on a result of the comparison.

26. The system of claim 25, wherein the one or more devices are further configured to repeat the selecting a candidate, estimating a weight, forming a new rule, comparing a likelihood of the data, and selectively adding the new rule for a number of iterations.

27. The system of claim 25, wherein the one or more devices are further configured to repeat the estimating a weight, forming a new rule, and comparing a likelihood of the data for a number of iterations.

28. The system of claim 25, wherein the training data is associated with a plurality of prior searches and the features include at least one of user information corresponding to users who provided search queries, query data corresponding to the search queries, and document information corresponding to documents relating to the search queries.

29. The system of claim 25, wherein when selecting a candidate condition, the one or more devices are configured to at least one of:

create the candidate condition from combinations of the features,

create the candidate condition from a single one of the features, and

create the candidate condition from a complement of at least one of the features.

30. The system of claim 25, wherein the features include at least one million features.

31. The system of claim 25, wherein when estimating a weight, the one or more devices are configured to determine a weight that maximizes a likelihood of the training data given the model.

32. The system of claim 25, wherein when selectively adding the new rule to the model, the one or more devices are configured to add the new rule to the model when the likelihood of the training data when the model includes the new rule is sufficiently greater than when the model does not include the new rule.

33. The system of claim 25, wherein when selectively adding the new rule to the model, the one or more devices are further configured to:

associate a cost with the candidate condition, and

determine whether to add the new rule to the model based, at least in part, on the cost associated with the candidate condition.

34. The system of claim 25, wherein the one or more devices include a plurality of devices, one of the plurality of devices being configured to:

select the candidate condition,

request information associated with the candidate condition from other ones of the devices,

receive the requested information from the other devices, and

estimate the weight for the candidate condition based, at least in part, on the requested information.

35. The system of claim 34, wherein the requested information includes predicted probabilities associated with the candidate condition.

36. The system of claim 34, wherein the one device is further configured to transmit information regarding the candidate condition and the estimated weight to the other devices.

37. The system of claim 34, wherein the training data further includes a plurality of instances, each of the features corresponding to one or more of the instances; and

wherein each of the devices is responsible for a subset of the instances and includes a feature-to-instance index that maps features to instances to which the features correspond.

38. A method for generating a model, comprising:

selecting candidate conditions from training data;

estimating weights for the candidate conditions;

forming new rules from the candidate conditions and corresponding ones of the weights;

comparing a likelihood of the training data between a model with the new rules and the model without the new rules; and

selectively adding the new rules to the model based, at least in part, on results of the comparing.

39. A system for generating a ranking model, comprising:

a repository configured to store data corresponding to a plurality of prior searches; and

one or more devices configured to:

select a candidate condition based, at least in part, on the data in the repository,

estimate a weight for the candidate condition,

form a new rule from the candidate condition and the weight,

compare a likelihood of the data between a model with the new rule and the model without the new rule, and

selectively add the new rule to the model based, at least in part, on a result of the comparison.

40. A method for ranking documents, comprising:

receiving a search query;

identifying documents relating to the search query;

determining prior probabilities of selecting each of the documents;

determining a score for each of the documents based, at least in part, on the prior probability of selecting the document; and

generating search results for the search query from the scored documents.

41. The method of claim 40, wherein the prior probability of selecting one of the documents is determined based, at least in part, on data regarding at least one of a position of the document within search results, a prior score assigned to the document, and a number of documents above the document in the search results that were selected.

42. A system for generating a model, comprising:

a repository configured to store training data that includes a plurality of instances and a plurality of features, each of the features corresponding to one or more of the instances; and

a plurality of devices, at least one of the devices being configured to:

create a feature-to-instance index that maps features to instances to which the features correspond,

select a candidate condition,

request information associated with the candidate condition from other ones of the devices,

receive the requested information from the other devices, estimate a weight for the candidate condition based, at least in part, on the requested information,

form a new rule from the candidate condition and the weight, and

selectively add the new rule to the model.

43. The system of claim 42, wherein the at least one of the devices is further configured to transmit information regarding the new rule to the other devices.

44. A system for generating a model, comprising:

a repository configured to store a plurality of instances; and

a plurality of devices, at least one of the devices being configured to:

analyze a subset of the instances to identify matching candidate conditions,

analyze the candidate conditions to collect statistics regarding predicted probabilities from the matching instances,

gather statistics regarding one of the candidate conditions from other ones of the devices,

determine a weight associated with the one candidate condition based on at least one of the collected statistics and the gathered statistics,

form a rule from the one candidate condition and the weight, and

selectively add the rule to the model.

45. The system of claim 44, wherein the at least one of the devices is further configured to record information concerning the instances and the matching candidate conditions as condition-instance pairs.

46. The system of claim 44, wherein the at least one of the devices is further configured to output the rule to the other devices when the rule is added to the model.